

IE and QA

wrap-up

Guy Lapalme
Université de Montréal

What did we learn?

- language is the way of showing our intelligence
- NLP help us create intelligent machines which will help humans
- understanding language is fundamental
 - first works on *deep* understanding in artificial situations did not scale up
 - recent work focus on *real text* with more limited goals

Information Extraction

- Extract structured information from unstructured texts
 - named entities
 - attributes of entities
 - relations between entities
- Bottom-up approaches
 - based on patterns developed (or learned) in specific domains
- MUC conferences showed the importance of evaluation

Question-Answering

- Find *the answer* and nothing but the answer
- A common architecture has evolved
 - Question analysis
 - Document preprocessing
 - Answer generation
- Answer must often combine several information sources
- TREC-QA evaluations established the domain

Overall

- IE and QA are *instances* of the more general goal of NLP techniques
 - information retrieval
 - summarization
 - text mining
- Frontiers between them are fuzzy
- Many applications
 - e-commerce (customer relations)
 - semantic web